



Linx

Revue des linguistes de l'université Paris X Nanterre

78 | 2019

La linguistique des genres, en actes et en questions

Oral représenté et narration en ancien français. Spécificités syntaxiques dans trois textes de genres distincts

Represented oral speech and narration in Old French: Syntactic specificities in three texts of different genres

Nicolas Mazziotta et Julie Glikman



Édition électronique

URL : <http://journals.openedition.org/linx/3151>

DOI : 10.4000/linx.3151

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Référence électronique

Nicolas Mazziotta et Julie Glikman, « Oral représenté et narration en ancien français. Spécificités syntaxiques dans trois textes de genres distincts », *Linx* [En ligne], 78 | 2019, document 6, mis en ligne le 30 juin 2019, consulté le 25 juillet 2019. URL : <http://journals.openedition.org/linx/3151> ; DOI : 10.4000/linx.3151

Ce document a été généré automatiquement le 25 juillet 2019.

Département de Sciences du langage, Université Paris Ouest

Oral représenté et narration en ancien français. Spécificités syntaxiques dans trois textes de genres distincts

Represented oral speech and narration in Old French: Syntactic specificities in three texts of different genres

Nicolas Mazziotta et Julie Glikman

1. Introduction

- 1 Cette contribution a pour objet l'étude des spécificités syntaxiques de l'oral représenté et de la narration en français médiéval. Elle s'appuie sur les études ayant démontré la pertinence d'une telle distinction pour l'analyse syntaxique (→1.1). En employant des techniques statistiques exploratoires (analyses factorielles des correspondances multiples), nous étudierons la stabilité des spécificités repérées antérieurement de la distinction entre oral représenté et narration sur le Graal (Glikman & Mazziotta 2013) dans deux textes très éloignés typologiquement : la *Chanson de Roland* et la *Conquête de Constantinople* de Robert de Clari, qui seront comparés à la *Queste del saint Graal* (→1.2). Nous cherchons ainsi à voir si la distinction oral représenté / narration est stable à travers les différents types de textes, ou si d'autres paramètres sont à prendre en compte.

1.1. Résultats précédents

- 2 La dichotomie désormais classique qui oppose oralité et scripturalité (Koch & Oesterreicher 2001) a amené les chercheurs travaillant sur l'histoire de la langue française à réfléchir sur l'intérêt de cette distinction pour des variétés qui ne nous ont été transmises que par voie écrite. Les études récentes ont contribué à développer le concept d'*oral représenté* (désormais « OR » ; Marchello-Nizia 2012) pour rendre compte de la

manière dont des faits oraux sont transmis par écrit dans les textes qui nous sont parvenus, au travers du « discours direct » (Rodriguez Somolinos 2013, Lagorgette & Larrivée 2013, Rodriguez Somolinos 2016, Ayres Bennet et al. 2018). Ce dernier est reconnaissable par des indices propres à la situation d'énonciation : emploi des pronoms nominaux de la première et de la deuxième personne, impératifs et vocatifs, interjections. Ces indices définissent *a priori* l'opposition entre deux types discursifs : l'OR et la narration « normale » (désigné ci-dessous par le terme *narration*).

- 3 Selon cette conception, les études prennent en compte le caractère artificiel de ce qui n'est en fin de compte que l'image déformée de l'oralité tout en se focalisant sur la recherche de tendances spécifiquement orales. Pareil angle d'attaque permet notamment de réévaluer la caractérisation de certains textes. Ainsi, les *Dialogues du pape Grégoire*¹, bien qu'entièrement dialogué, a des caractéristiques proches de la narration (Guillot et al. 2015). Le discours direct nous donne un accès *indirect et artificiel* à une facette authentique de la langue, que l'on peut interroger notamment sous l'angle de la pragmatique historique (Lagorgette et al. 2006 ; Opermann-Marsaux & Rodríguez Somolinos 2008 ; Rodríguez Somolinos 2013, 2016) ou sous celui de la morphosyntaxe historique. C'est cette dernière approche que nous privilégions ici. La recherche sur les caractéristiques grammaticales de l'OR a mis en évidence son caractère spécifique par rapport à la narration, qu'il s'agisse des temps verbaux (Marchello-Nizia 2012), des parties du discours² ou des fonctions syntaxiques mobilisées³. Les premières études ont ainsi démontré l'existence de spécificités relativement stables en diachronie (Guillot et al. 2013, 2015, 2017b), proches de l'évolution connue de la langue (Glikman & Mazziotta 2013). Ces conclusions nous mènent raisonnablement à penser que l'OR a bel et bien des caractéristiques de l'oralité véritable. Glikman & Mazziotta 2013 mettaient en évidence que l'OR se caractérise par un emploi plus fréquent d'un syntagme ayant la fonction de sujet, par l'emploi de formes analytiques du verbe actif (auxiliaire et participe passé ou « verbe modal » accompagné d'un infinitif) et par une fréquence plus élevée de particules de négation (*pas, point, etc.*) – sans que cette fréquence plus élevée ne soit clairement caractéristique de l'OR. Ces caractéristiques rapprochent typologiquement l'OR du français moderne. Ces observations sont en accord avec la plupart de celles de Guillot et al. 2013, qui remarquaient en ce qui concerne les parties du discours que les pronoms et les infinitifs après des verbes modaux étaient surreprésentés en OR. Ils faisaient par contre remarquer que les participes passés sont plutôt spécifiques de la narration, ce qui paraît contradictoire avec les observations de Glikman & Mazziotta 2013. Par opposition, selon les mêmes études, la narration se caractérise principalement par l'emploi de mots-outils articulant les énoncés entre eux (*et, mais* initiaux, etc.).
- 4 Par *spécificité* il faut entendre ici « tendance » ; p.ex. : les phrases de l'OR ont tendance à comporter davantage de sujets exprimés par un terme pronominal, mais il est fréquent que les phrases de la narration en comportent également. En dehors des critères définitoires *a priori* que nous mentionnons plus haut, l'opposition entre les deux types discursifs n'est pas catégorique. Il est donc difficile de l'évaluer intuitivement.
- 5 Dans cet article, à la suite de Glikman & Mazziotta 2013, nous nous intéressons de manière fine à l'analyse des fonctions syntaxiques. En cela, nos travaux sont complémentaires des travaux de Guillot et al. 2013, 2015 et 2017b, qui sont basés sur l'analyse des parties du discours.

1.2. Questions et corpus

- 6 On connaît les difficultés qu'il y a à démêler les paramètres de variation : les données langagières sont affectées par le temps, l'espace, la société et la situation d'énonciation (axes de la *diavariation* ; Glessgen 2007 : ch. 1.5). Nous voudrions ici évaluer la stabilité morphosyntaxique de l'oral représenté, observée jusqu'ici de manière globale, en comparant deux textes volontairement sélectionnés pour leur éloignement historique, géographique et typologique (incluant la notion de *genre textuel*) : la *Chanson de Roland* (ci-après « roland ») et la *Conquête de Constantinople* de Robert de Clari (ci-après « clari »). Les caractéristiques de ces textes seront contrastées, à titre de comparaison, avec celles de *La quête del saint Graal* (ci-après « qgraal »).
- *clari*⁴ est postérieur à 1205 et rédigé en scripta⁵ picarde. Il s'agit d'une chronique en prose⁶ rédigée par un chevalier parti à la croisade. Les caractéristiques syntaxiques et stylistiques du texte sont étudiées par Dembowski (1963).
 - *roland*⁷ est daté du deuxième quart du 12^e siècle (pour la datation du manuscrit, la date de composition étant estimé aux environs de 1100 (BFM)), et rédigé en scripta anglo-normande. Il s'agit d'une chanson de geste⁸. L'étude stylistique de Rychner (1955) a mis en évidence l'importance des formules dans la structure de la rédaction.
 - *qgraal*⁹ est daté de 1225-1230 et rédigé dans une scripta peu marquée dialectalement. Il s'agit d'un roman¹⁰ arthurien en prose. C'est ce texte dont nous avons étudié les 1007 premières phrases dans Glikman & Mazziotta 2013¹¹ et pour lequel les tendances précitées ont été mises en évidence (→1.1).
- 7 La question fondamentale à laquelle nous essayons de répondre ici est celle de la stabilité de la distinction entre l'OR et la narration. Pour investiguer, nous avons choisi de faire usage de méthodes statistiques exploratoires (heuristiques), qui nous permettent de mettre en évidence les associations les plus fortes entre les types discursifs et les caractéristiques morphosyntaxiques examinées.

2. Analyse

- 8 Nous présenterons ici les variables par lesquelles les données sont encodées (→2.1) avant de les soumettre globalement à une analyse factorielle des correspondances multiples (→2.2). Pour des raisons dont la pertinence apparaîtra au fil des sections qui suivent, nous traiterons ensuite séparément les phrases et les propositions subordonnées (→2.3), puis nous examinerons les textes séparément (→2.4).

2.1. Préparation des profils des individus examinés

- 9 Les textes étudiés sont traités sur la base de leur numérisation dans la *Base de français médiéval* (Guillot et al. 2017a, désormais « BFM »). Le corpus est enrichi par trois couches d'annotations distinctes : (i) une identification des parties du discours (CATTEX2009 ; Prévost et al. 2013) élaborée par l'équipe de la BFM ; (ii) une identification du discours direct, également élaborée par cette équipe ; (iii) une annotation syntaxique exhaustive élaborée par l'équipe du *Syntactic Corpus of Medieval French* (désormais « SRCMF » ; Prévost et Stein 2013 ; Mazziotta 2012). Cette dernière couche mérite une brève introduction : elle suit un modèle d'analyse dépendancielle, i.e. où les mots sont groupés par des relations

hiérarchisées et typées. Par exemple, dans *Il morout*, *Il* est groupé à *morout* par une relation de type « sujet », qui est orientée de *morout* à *il*. Le premier terme est dit « gouverneur » du second, qui est dit « dépendant ». Cette analyse peut être représentée graphiquement de diverses manières ; p.ex. par le diagramme de la fig. 1, qui représente l'analyse de (1), où les relations sont représentées par des flèches étiquetées en petites capitales.

1. Des lors te toli li anemis la veue. 'Dès cet instant, le Diable t'a dépouillé de ta vue.' (*qgraal*, 190a, 1)

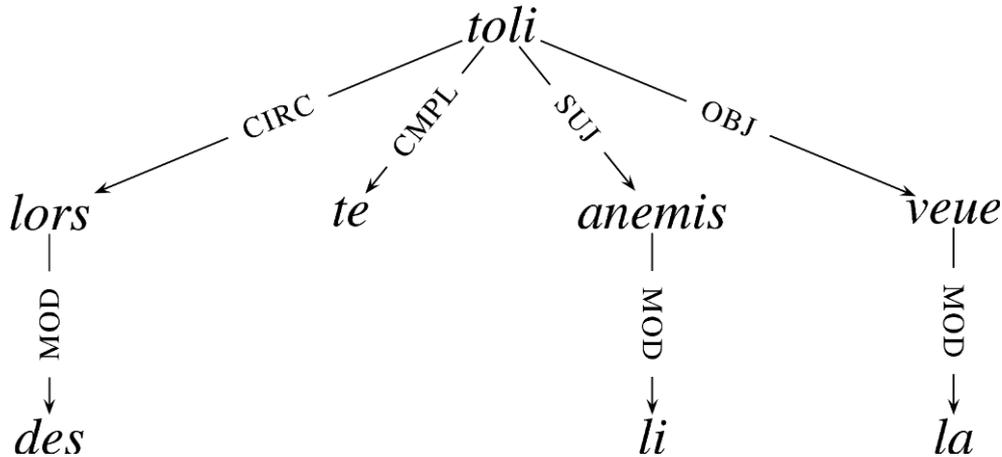


Figure 1. Diagramme de l'exemple (1)

- 10 Les individus statistiques que nous sélectionnons sont des propositions à verbe fini, i.e. des structures syntaxiques dont le terme gouverneur principal est un verbe fini (pour les langues romanes, il s'agit d'un verbe conjugué à un mode « personnel »). Chaque proposition est décrite à l'aide des dix variables sélectionnées principalement pour leur pertinence remarquée dans les études précédentes.

1. Type discursif et texte. – Six modalités (i.e. « valeurs de variable ») combinant le type discursif (OR pour « oral représenté » et N pour « narration ») et le nom du texte ; p.ex. : OR-clari ou N-roland¹².
2. Type de proposition. – Huit modalités correspondant à la dépendance syntaxique de la proposition¹³ : Phrase pour les indépendantes (Bischof et al. 2012 : § Snt), SujetPersonnel pour les subordonnées sujets (*ibid.* : § SjPer), Objet pour les complétives objet (*ibid.* : § Obj), Regime pour les interrogations indirectes (*ibid.* : § Reg, 5), Complément pour les autres subordonnées non circonstancielles dépendant du verbe (*ibid.* : § Cmpl, 5), Circonstant pour les subordonnées circonstancielles (*ibid.* : § Circ), Relative pour les propositions introduites par un pronom relatif (*ibid.* : § ModA), Determinative pour les propositions non relatives dépendant d'un gouverneur non verbal (*ibid.* : § ModA).
3. Sujet (Sj). – six modalités décrivant le sujet (Bischof et al. 2012 : § Sj) : SJ_0 pour « pas de sujet exprimé en dehors du verbe », SJ_NOM pour « sujet nominal » et SJ_PRO pour « sujet pronominal personnel », SJ_IMP pour « sujet impersonnel » (Bischof et al. 2012 : § SjImp) et SJ_AUTRE dans tous les autres cas.
4. Objet (OBJ). – Cinq modalités décrivant l'objet (Bischof et al. 2012 : § Obj) : OBJ_0, OBJ_NOM, OBJ_PRO, OBJ_AUTRE suivant le même principe que pour les modalités de la variable précédente, mais aussi Obj_VFin « avec subordonnée objet ».
5. Relateur (REL). – Deux modalités décrivant la présence ou l'absence de la fonction relateur (préposition ou conjonction ; Bischof et al. 2012 : § RelNC) : REL_0 pour « absence de relateur », REL_1 « présence d'un relateur ». En pratique, cette variable décrit principalement la présence de *et*, *mais* ou *ains* 'au contraire, par contre' en début de phrase.
6. Particule de négation de type *pas*, *point*, *mie*, etc. (NGPRT). – Deux modalités décrivant la présence ou l'absence : NGPRT_0 ou NGPRT_1.

7. Attribut du sujet (ATSJ). – Deux modalités décrivant la présence ou l’absence : *ATSJ_0* ou *ATSJ_1*.

11 Les deux variables suivantes décrivent la présence de formes verbales analytiques :

1. *AUXA* pour « forme au participe passé employée avec un auxiliaire pour construire un temps actif » (Bischof *et al.* 2012 : § AuxA). – Deux modalités décrivant la présence ou l’absence : *AUXA_1* ou *AUXA_0*.

2. *AUXI* pour « forme à l’infinitif introduite par un « auxiliaire modal » (vouloir, devoir, pouvoir, savoir, soloir ; Bischof *et al.* 2012 : *ibid.*). – Deux modalités décrivant la présence ou l’absence : *AUXI_1* ou *AUXI_0*.

12 Enfin, la dernière variable évalue la présence de l’adverbe *si* “de cette manière” dans la proposition, en vertu de son rôle bien connu d’articulateur d’énoncés (Marchello-Nizia 1985) :

3. Variable *SI*. – Deux modalités décrivant la présence ou l’absence : *SI_1* ou *SI_0*.

13 Cela nous donne 13 373 individus (un individu correspondant à une proposition), qui fournissent 1 987 profils attestés sur 92 160 profils possibles compte tenu des variables énoncées ci-dessus¹⁴, ce qui nous permet d’anticiper la présence d’associations très fortes et conventionnalisées. Sur les 1 987 profils attestés, les 9 profils du tableau 1 ci-dessous sont les seuls à représenter au moins 1 % des individus :

Tableau 1. Profils représentant plus de 1% des individus

T. Disc.	REL	T. Prop	SJ	OBJ	REL	NGPRT	ATSJ	AUXA	AUXI	SI	Fq. abs.	Fq. rel.
N-roland	0	Phrase	NOM	0	0	NA	0	0	0	0	386	2,89
N-roland	0	Phrase	0	NOM	0	NA	NA	0	0	0	304	2,27
N-graal	1	Phrase	0	0	1	NA	0	0	0	0	257	1,92
N-roland	0	Phrase	0	0	0	NA	0	0	0	0	239	1,79
N-clari	0	Phrase	0	0	0	NA	0	0	0	1	221	1,65
N-clari	NA	Relative	NA	NA	NA	NA	0	0	0	0	204	1,53
N-qgraal	NA	Relative	NA	NA	NA	NA	0	0	0	0	168	1,26
N-qgraal	1	Phrase	0	NOM	1	NA	NA	0	0	0	134	1,00

14 De nombreuses variables morphosyntaxiques ont été ignorées en raison de leur trop faible fréquence (p.ex. les attributs de l’objet), ou de leur caractère non contraint (compléments circonstanciels divers, structures coordonnées). Par ailleurs, les interférences évidentes ont également été neutralisées dans le traitement statistique¹⁵.

2.2. Analyse globale

15 Pour évaluer les relations entre les différentes modalités des variables observées, nous nous servons de l’analyse factorielle des correspondances multiples (désormais "ACM" ; Lebart & Salem 1994, Lebart *et al.* 1997). Il s’agit d’une technique statistique heuristique qui met en évidence, notamment par des représentations graphiques du type de fig. 2, les tendances d’attraction et de répulsion de modalités catégorielles. Contrairement aux tests de distribution, cette technique ne statue pas sur la significativité des attractions, mais

elle permet de les hiérarchiser. Pour ce faire, elle décompose l'information contenue dans de grandes tables dont les lignes ressemblent à celles de tab.1. Le résultat de la décomposition de cette table complexe est un ensemble de vecteurs simples et monodimensionnels qui comportent une partie de l'information (appelée *inertie*). Ces vecteurs peuvent être représentés graphiquement comme des coordonnées sur un espace dont le nombre de dimensions équivaut au nombre de vecteurs conservés (deux dans fig.2 et les graphiques qui suivent). Dans les graphiques, nous avons représenté les individus par des points grisés et les modalités par des traits étiquetés de leur nom et de leur fréquence absolue. Les modalités contribuent de manière plus ou moins forte à l'information contenue dans chaque vecteur. Sur les figures qui suivent, nous avons uniquement conservé les modalités dont la contribution était suffisamment importante (supérieure à la moyenne). Nous avons représenté par < — > celles qui contribuent le plus au premier axe et par < | > celles qui contribuent au second axe ; les deux symboles se combinent en une croix quand la modalité contribue aux deux axes. La proximité des modalités et l'angle qu'elles forment avec la coordonnée centrale permettent d'évaluer les attractions : un angle aigu est le signe d'une attraction ; un angle obtus est le signe d'une répulsion.

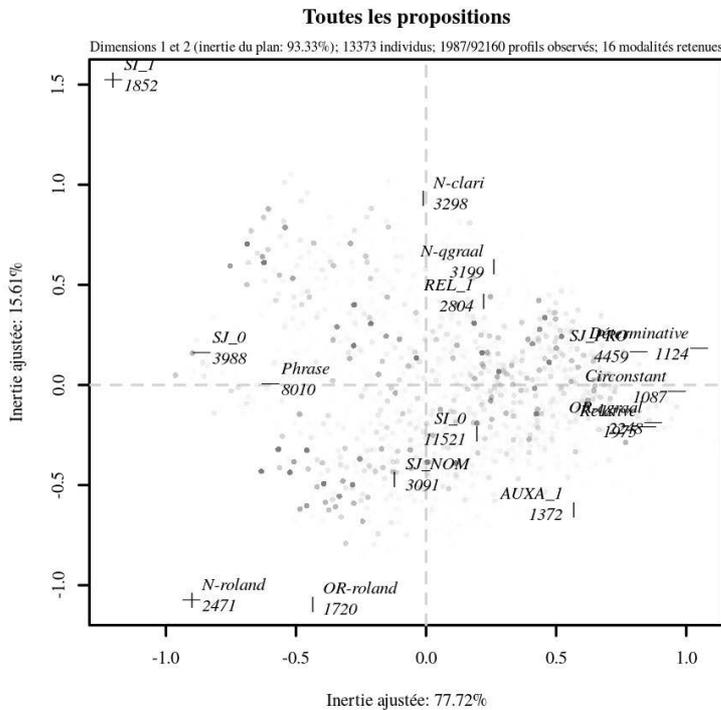


Figure 2. Toutes les propositions

- 16 À première vue, l'analyse globale de tous les individus montre (axes 1 et 2 ; fig. 2) que les types discursifs/textes sont groupés en trois ensembles : roland (OR ou narration) se trouve en bas à gauche ; la narration des autres textes se trouve au-dessus ; l'oral OR dans qgraal se trouve à droite. N-roland se caractérise par une préférence pour les indépendantes plutôt que les subordonnées et l'absence de sujet ou le sujet nominal (on voit que les points N-roland forment un angle aigu avec phrase et SJ_0 d'une part et avec SJ_NOM d'autre part). En cela, il contraste très fort avec l'OR dans qgraal, qui se démarque par l'emploi de subordonnées et de sujets pronominaux (les modalités sont très proches

et forment des angles très aigus). La narration dans les deux textes en prose est en outre caractérisée par l'emploi d'outils grammaticaux permettant l'articulation du texte : *si* et, pour ce qui concerne le contraste avec *roland*, les relateurs. Il est frappant que les variables discursives et les textes apparaissent immédiatement parmi les paramètres les plus structurants du nuage de données, mais que l'opposition entre OR et narration est moins structurante dans *roland* que dans *qgraal*, puisqu'il faut observer le second plan (axes 2 et 3 ; fig. 3) pour la voir clairement apparaître et que le troisième axe exprime une part très faible de l'information (moins de 5 %) : on voit que la narration de *roland* favorise les sujets nominaux et les subordonnées relatives et circonstancielles (soit une certaine forme de complexité syntaxique). Une association entre la narration et les formes analytiques actives avec participe passé (*AUXA_1*) s'oppose en outre à une association entre l'OR et les formes analytiques actives avec infinitif (association difficile à analyser) et les sujets consistant en des infinitifs.

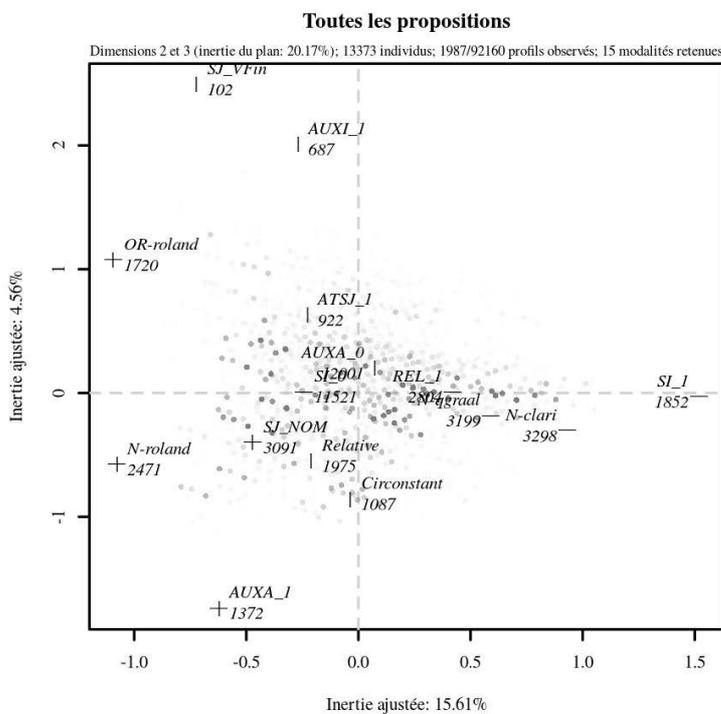


Figure 3. Toutes les propositions-2

- 17 Quoi qu'il en soit, cette première ACM montre l'importance, d'une part, de la distinction entre les indépendantes et les subordonnées et, d'autre part, de la répartition différente selon les textes des caractéristiques qui opposent OR et narration. Il est donc raisonnable de traiter de manière séparée les différents textes et les différents types de propositions.

2.3. Analyse par type de proposition

2.3.1. Les propositions principales et autonomes

- 18 Les deux premiers axes de l'ACM portant sur les phrases uniquement (propositions principales ou autonomes) (fig. 4 ; 92,27 % de l'information) montrent que les phrases de *roland*, quel que soit le type discursif, s'opposent principalement à celles de la narration dans les textes plus récents en prose.

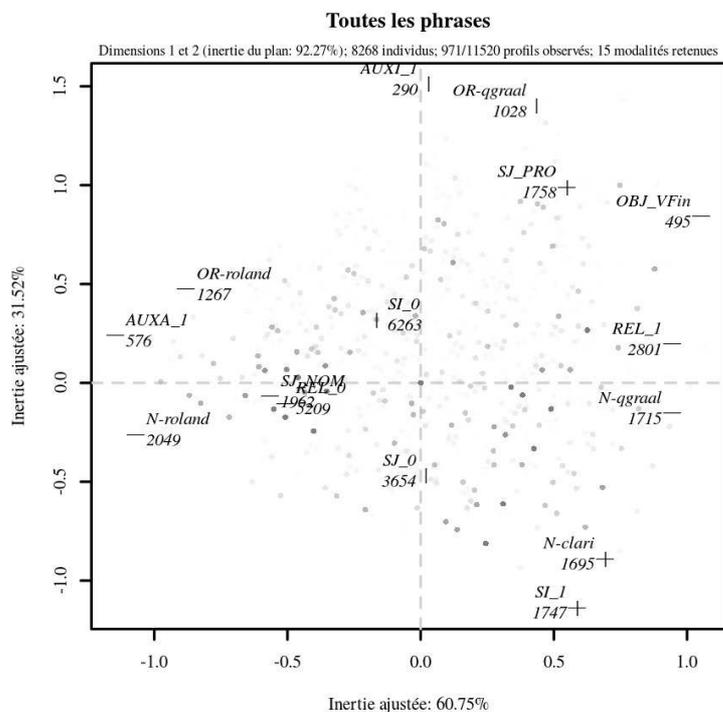


Figure 4. Toutes les propositions principales et indépendantes ('phrases')

- 19 Ainsi *roland* présente davantage de sujets nominaux (contre des sujets pronominaux pour la narration des autres textes) et se caractérise en outre par une tendance à ne pas utiliser de relateurs (les autres textes emploient davantage des relateurs ou *si*). On observe en outre que *roland* s'oppose à l'emploi de complétives objets et comporte globalement plus de formes verbales analytiques actives, ce qui l'oppose à ce qui avait été observé pour la narration de *qgtaal* (→1.1) et réapparaît sur le plan. L'observation des données révèle que *roland* n'emploie pratiquement que le passé composé dans ses indépendantes (2).

(2) *Perdud avuns Espagne, nostre tere 'nous avons perdu l'Espagne, notre terre' (roland, p. 160, v. 2119)*

- 20 L'OR de *qgtaal* (sujets pronominaux, tours périphrastiques avec infinitif, absence de *si*) s'oppose le plus fortement à la narration de *clari* (sujets nuls, emploi de *si*).

2.3.2. Les propositions subordonnées

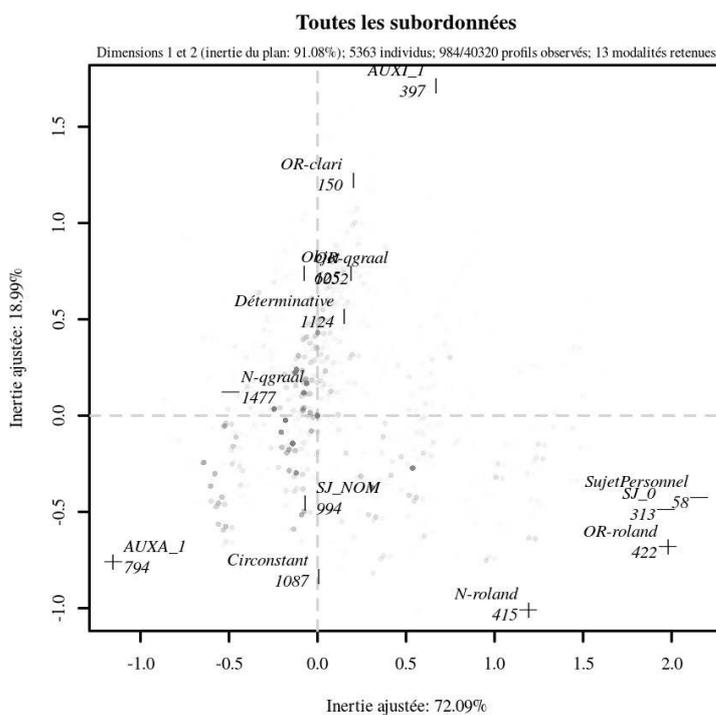


Figure 5. Toutes les subordonnées

- 21 Le premier axe de l'ACM effectuée sur les subordonnées (fig. 5 ; 91,08 % de l'information) montre que *roland* se distingue principalement de la narration de *qqraal* par une prédominance des subordonnées sans sujet et de celles qui assument la fonction sujet, comme dans (3)¹⁶.

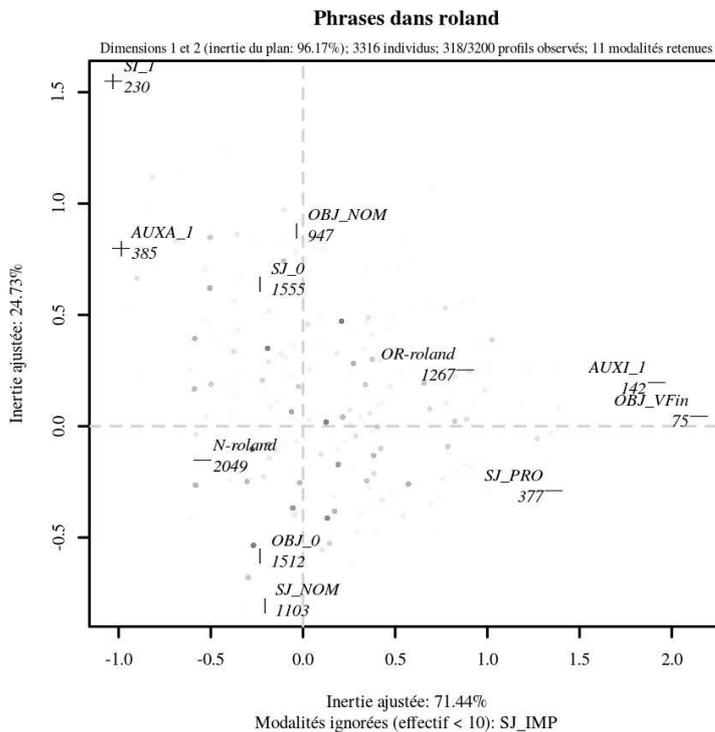
(3) Asez est melz qu'il i perdent lé chefs 'Ce serait bien mieux qu'ils y perdent la tête' (*roland*, v. 44)

- 22 L'identification d'une subordonnée sujet dans ce genre de cas est donc une particularité de SRCMF et ces structures pourraient être rapprochées de celles où le sujet n'est pas exprimé (ce qui est parfaitement conforme à leur position sur le plan). Cette caractéristique oppose fortement *roland* à la narration dans *qqraal*.
- 23 Selon le second axe, les subordonnées de l'OR dans *qqraal* et dans *clari* se caractérisent quant à elles par : l'emploi d'« auxiliaires modaux » + infinitif (*AUXI_1*), le rejet des sujets nominaux (*SJ_NOM*), la fonction de circonstant et les formes auxiliées actives accompagnées d'un participe passé (*AUXA_1*). Il apparaît donc que ces tendances quant à l'emploi de participes passés sont clairement différentes des observations précédentes sur les 1007 premières propositions indépendantes de *qqraal* (→1.1) et rejoignent celles de Guillot *et al.* 2013.

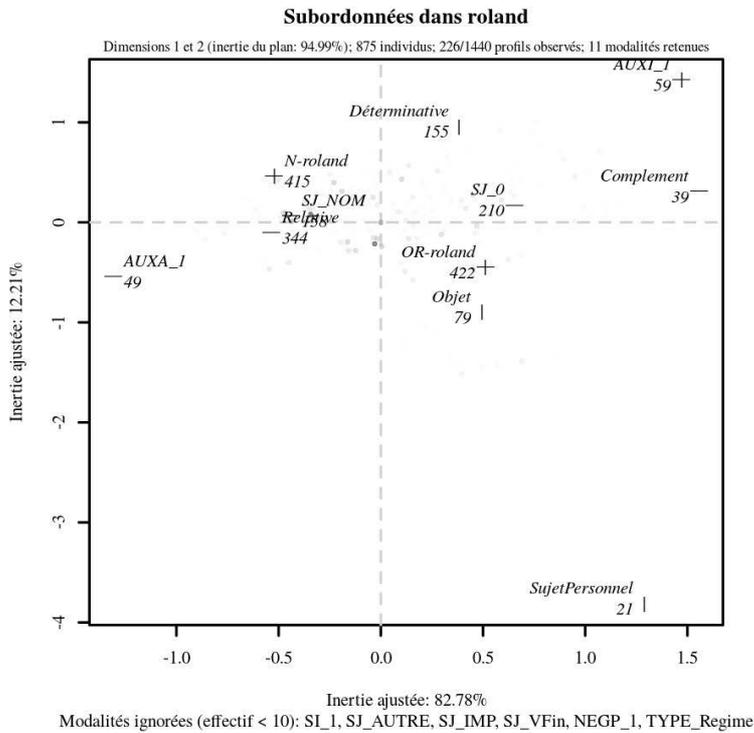
2.4. Analyse par texte : *Roland* et *Clari*

2.4.1. Roland

- 24 Les phrases de *roland* s'opposent effectivement suivant le type de discours (fig. 6, axe 1) : la narration se caractérise par l'emploi de formes auxiliées actives et de l'adverbe *si*, alors que l'OR se caractérise par l'emploi de formes infinitives auxiliées, d'objets infinitifs et de sujets pronominaux. L'opposition entre le sujet nominal et l'absence de sujet ne semble pas être liée au type de discours (axe 2).



- 25 Figure 6. Les propositions principales et indépendantes dans *Roland*
- 26 En ce qui concerne les subordonnées, les deux types discursifs s'opposent également : l'OR a plus tendance à employer des sujets nuls et infinitifs auxiliés, des complétives objets, alors que la narration emploie davantage de relatives et de déterminatives, de sujets nominaux et de constructions auxiliées actives. Les formes périphrastiques infinitives ne paraissent pas liées à l'un ou à l'autre type discursif.



27 Figure 7. Les subordonnées dans *Roland*

2.4.2. Clari

28 Dans les phrases, c'est l'OR, plus rare, qui attire la spécificité de la présence de sujets pronominaux et de formes auxiliées actives, exactement comme dans *qgraal* pour ce qui concerne les principales et indépendantes (Fig. 4). Indépendamment du type de discours, on constate la concurrence entre les relateurs (*et, mais, etc.*) et *si*.

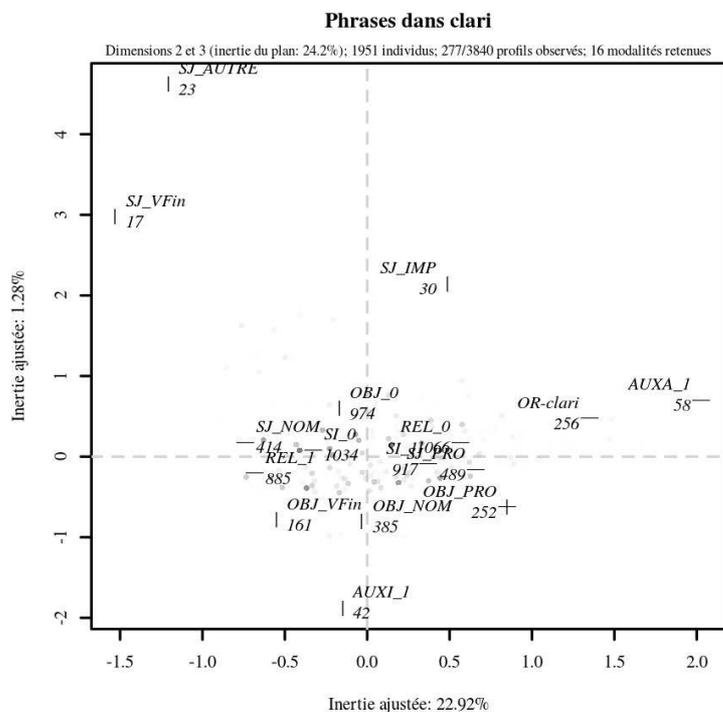


Figure 8. Les propositions principales et indépendantes dans Clari

- 29 Dans les subordinées en OR, on retrouve la même attraction pour les formes périphrastiques infinitives et une opposition très nette avec l'emploi de *si*, les participes passés auxiliés (*AUXA_I*) et les circonstanciels. On remarque par ailleurs que les structures circonstanciels et les tours périphrastiques actifs sont proches des circonstants et éloignés de *si*. Un retour aux données montre que cela correspond à l'emploi de formules littéraires du type de (4), souvent présentées comme caractéristiques de la prose médiévale (Marchello-Nizia 1985) et ici typique de la narration (du « non-OR »).

(4) Quant on eut mengié, si se departirent li baron [...] (*clari*, p. 97)

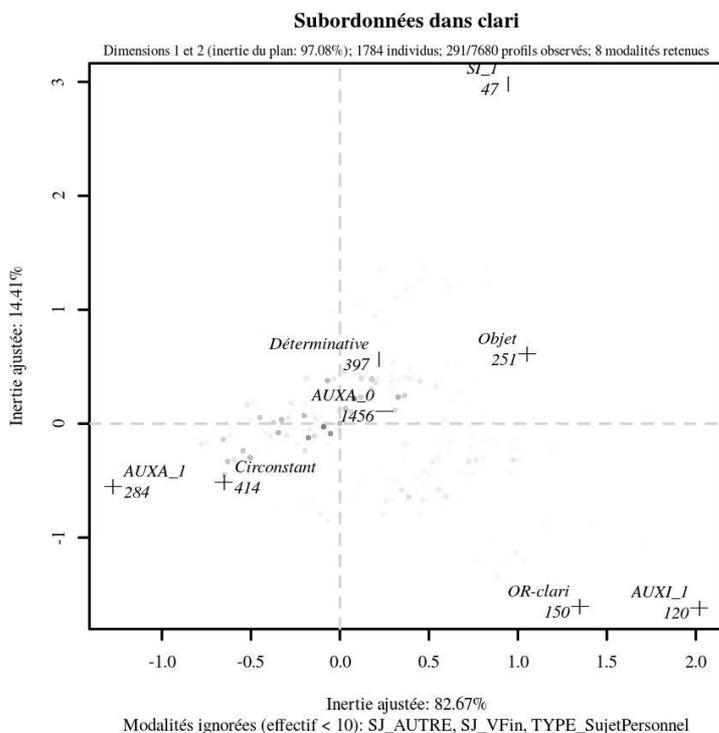


Figure 9. Les subordonnées dans Clari

3. Conclusion

- 30 Les analyses ont montré que la principale partition est celle qui oppose les propositions indépendantes (→2.2) aux subordonnées et *roland* aux autres textes. La prééminence d'emploi de phrases autonomes, non articulées aux autres par des mots outils spécifiques (*et, si, etc.*), dans *roland* caractérise aussi bien l'OR que la narration. Le texte comporte également de nombreuses phrases employant le passé composé. Le style narratif est ainsi conçu de manière générale comme moins « articulé » (Rychner 1955), avec moins de constructions par subordination et très peu de temps composés autres que le passé composé.
- 31 Toutefois, l'OR de *roland* partage des caractéristiques avec celui de *qgraal* et de *clari* : l'expression pronominale du sujet des phrases et l'emploi de formes infinitives avec auxiliaire modal paraissent des caractéristiques stables de l'OR, alors que l'emploi de *si* ou d'un autre articulateur (en concurrence) est plutôt caractéristique de la narration. En particulier, *clari* emploie des tours formulaïques en *quand... si...* (→2.4.2) qui sont spécifiques de la narration, ce qui nous conforte dans l'idée que le texte est produit par imitation de modèles et par utilisation d'un nombre limité de procédés littéraires (Dembowski 1963). On remarquera que l'emploi des particules de négation n'est pas du tout saillant.
- 32 Les études précédentes avaient mis en évidence une apparente contradiction en ce qui concerne l'emploi des participes passés : typiques de l'OR pour Glikman & Mazziotta 2013, au travers de l'emploi des temps composés, et typiques de la narration pour Guillot *et al.* 2013. L'examen séparé selon les textes et selon les différentes sortes de propositions a permis de résoudre ce problème : d'une part, il apparaît que *roland* dans son ensemble a

plutôt tendance à employer le passé composé et, d'autre part, il apparaît que les temps composés sont également plus fréquents dans les subordonnées.

- 33 De manière plus générale, la présence d'une opposition OR/narration ne fait aucun doute pour les textes étudiés, mais il ne s'agit pas du seul paramètre influant sur la syntaxe. Malgré une indéniable stabilité des caractéristiques de l'OR et de la narration dans des textes aussi différents que ceux dont nous avons parlé, notre étude montre également qu'il est nécessaire de distinguer l'analyse des subordonnées de celle des principales. Le type de texte doit aussi être pris en considération. L'opposition entre *roland* et les autres textes est plus forte que l'opposition entre oral représenté et narration. Reste encore à étudier si cette particularité de *roland* est liée au genre, à la forme (en vers tandis que les autres textes de l'étude sont en prose), ou encore à l'espace temporel (presque 2 siècles séparant ce texte des deux autres).

BIBLIOGRAPHIE

AYRES BENNETT, W., CARLIER, A., GLIKMAN, J., RAINSFORD, T. R., SIOUFFI, G., SKUPIEN DEKENS, C. (dir.), 2018, *Nouvelles voies d'accès au changement linguistique*, Classiques Garnier.

BISCHOF, B.-B. et al., 2012, *Guide d'annotation SRCMF*. Direction du projet : Sophie Prévost et Achim Stein. Conception technique de la documentation : Nicolas Mazziotta. Structure conceptuelle de la documentation : Julie Glikman et Nicolas Mazziotta. Rédaction de la documentation : Beatrice-Barbara Bischof et Julie Glikman.

DEMBOWSKI, P. F., 1963, *La Chronique de Robert de Clari : étude de la langue et du style*, University of Toronto Press.

Dictionnaire des Lettres Françaises, Le Moyen Age, 1964 (1992), Fayard.

GLESSGEN, M.-D. 2007. *Linguistique romane : domaines et méthodes en linguistique française et romane*. Paris : A. Colin.

GLIKMAN, J., MAZZIOTTA, N., 2013, « Représentation de l'oral et structures syntaxiques dans la prose de la *Queste del saint Graal* (1225-1230). » dans Lagorgette, D. et Larrivée P. (éds), *Représentations du sens linguistique 5*, Chambéry, Éditions de l'Université de Savoie, p. 69-90.

GUILLOT, C., HEIDEN, S., LAVRENTIEV, A., PINCEMIN, B., 2015, « L'oral représenté dans un corpus de français médiéval (9e-15e) : approche contrastive et outillée de la variation diasystémique. » dans *Les Variations diasystémiques et leurs interdépendances dans les langues romanes. Actes du Colloque DIA II à Copenhague (19-21 nov. 2012)*, p. 15-28, Éditions de linguistique et de philologie.

GUILLOT, C., LAVRENTIEV, A., PINCEMIN, B., HEIDEN, S., 2013 « Le discours direct au Moyen Age : vers une définition et une méthodologie d'analyse », dans Lagorgette, D. et Larrivée P. (éds) *Représentations du sens linguistique 5*. Chambéry, Éditions de l'Université de Savoie, p. 17-41.

GUILLOT-BARBANCE, C., HEIDEN, S., LAVRENTIEV, A., 2017a, « Base de français médiéval : une base de référence de sources médiévales ouverte et libre au service de la communauté scientifique », *Diachroniques*, n° 7, p.168-184.

- GUILLOT-BARBANCE, C., PINCEMIN, B., LAVRENTIEV, A., 2017b « Représentation de l'oral en français médiéval et genres textuels », *Langages* n° 208 (4/2017), p. 53-68, Armand Colin.
- KOCH, P., ÖSTERREICHER, W., 2001, « Gesprochene Sprache und geschriebene Sprache. Langage parlé et langage écrit », dans Holtus, G., Metzeltin, M., Schmitt, C. (éds), *Lexikon der romanistischen Linguistik*, Tübingen, Niemeyer, p. 584-627.
- LAGORGETTE, D., LARRIVÉE, P. (éds), 2013, *Représentations du sens linguistique 5*, Chambéry, Éditions de l'Université de Savoie.
- LAGORGETTE, D., OPERMANN-MARSAUX, E., RODRÍGUEZ SOMOLINOS, A. (éds), 2006, *Énonciation et pragmatique : approche diachronique, Langue française 149*.
- LEBART, L., SALEM, A., 1994, *Statistique textuelle*. Paris : Dunod.
- LEBART, L., MORINEAU, A., PIRON, M., 1997, *Statistique exploratoire multidimensionnelle*. Paris : Dunod.
- MARCHELLO-NIZIA, C., 1985, *Dire le vrai : l'adverbe « si » en français médiéval, essai de linguistique historique*, Droz, Genève.
- MARCHELLO-NIZIA, C., 2012, « L'oral représenté : un accès construit à une face cachée des langues 'mortes' » dans Guillot, C., Combettes, B., Lavrentiev, A., Oppermann-Marsaux, E., Prévost S. (éds) *Le Changement en français. Études de linguistique diachronique*, Bern/Berlin/Bruxelles, Peter Lang, 247-264.
- MAZZIOTTA, N., 2012, « Le Syntactic Reference Corpus of Medieval French : Structure, outils et exploitation. » dans Dister, A., Longrée, D., Purnelle, G. (éds.), *11^{es} Journées internationales d'Analyse statistique des Données Textuelles (JADT)*, p. 701-713. Liège, Bruxelles : Université de Liège.
- OPERMANN-MARSAUX, E., RODRÍGUEZ SOMOLINOS, A. (éds), 2008, *Propriétés énonciatives et pragmatiques du texte médiéval, L'Information grammaticale 118*.
- PRÉVOST S., GUILLOT C., LAVRENTIEV A., HEIDEN S., 2013, *Jeu d'étiquettes morphosyntaxiques CATTEX2009*, version 2.0, BFM - Base de Français Médiéval [En ligne]. Lyon : ENS de Lyon, < http://bfm.ens-lyon.fr/IMG/pdf/Cattex2009_2.0.pdf>.
- PRÉVOST S., STEIN, A., 2013, « Syntactic annotation of medieval texts: the Syntactic Reference Corpus of Medieval French (SRCMF). » dans P. Bennett, M. Durrell, S. Scheible & R. Whitt (eds.), *New Methods in Historical Corpora Corpus Linguistics and International Perspectives on Language, CLIP Vol. 3*, 275-282. Tübingen: Narr.
- REMACLE, L., 1948, *Le Problème de l'ancien wallon*, Liège, Bibliothèque de la Faculté de Philosophie et Lettres.
- RODRÍGUEZ SOMOLINOS, A. (éd), 2013, *Marques d'oralité en français médiéval, Diachroniques 3*.
- RODRÍGUEZ SOMOLINOS, A. (éd), 2016, *Énonciation et marques d'oralité dans l'évolution du français, Linx73*.
- RYCHNER, J., 1955, *La Chanson de geste : essai sur l'art épique des jongleurs* (Vol. 53). Droz, Genève.
- ZINK, M., 1992, *Littérature française du Moyen Age*, Paris, PUF.
- Corpus :
- Chanson de Roland*, édité par Segre C., Paris, Droz, 2003.
- Chanson de Roland*, édité par Gérard Moignet, Paris, Bordas, 1972. Publié en ligne par la Base de français médiéval, <http://catalog.bfm-corpus.org/roland>. Dernière révision le 2006-06-30

Robert de Clari, *Conquête de Constantinople*, édité par Dufournet J., Paris, Champion, 2004.

Robert de Clari, *Conquête de Constantinople*, édité par Lauer, P., Paris, Champion, 1924, BFM2013 (non disponible sur BFM2016)

Queste del saint Graal, édité par C. Marchello-Nizia et A. Lavrentiev, Lyon, Equipe BFM, 2013. Publié en ligne par la Base de français médiéval, http://catalog.bfm-corpus.org/qgraal_cm. Dernière révision le 2013-07-19.

NOTES

1. Tous les textes cités sont tirés du corpus de la *Base de français médiéval*, désormais BFM (Guillot et al. 2017). Voir →1.2.
2. Guillot et al. 2013, sur l'identification des parties du discours de près de deux millions de mots de la *Base de français médiéval*.
3. Glikman & Mazziotta 2013, sur l'analyse syntaxique fine de 1007 phrases de la *Queste del saint Graal*.
4. Éditions de référence : Dufournet 2004 ; Lauer 1924 pour la BFM.
5. On peut définir le terme *scripta* par « langue écrite régionale » : il ne s'agit pas d'un dialecte écrit, mais d'une variété écrite proche du français central qui comporte des traits dialectaux (Remacle 1948 ; pour une présentation de la notion, v. Glessgen 2007 : 343-345).
6. Alors que l'histoire s'écrit encore en latin, les chroniques en prose en français commencent à apparaître au début du 13^e s., voir Zink 2012.
7. Éditions de référence : Segre 2003 ; Moignet 1969 pour la BFM.
8. Poèmes épiques composés de laisses assonancées et chantés. Voir le *Dictionnaire des Lettres Françaises*, le Moyen Age, art. *chanson de geste*, p. 238. Roland est traditionnellement considéré comme la plus ancienne chanson de geste française.
9. Édition de référence : Marchello-Nizia/Lavrentiev 2013.
10. Le terme *roman* dans la littérature du moyen âge désigne un genre particulier, narratif et littéraire, d'abord en vers, puis en prose, qui se développe vers le 12^e s., soit après les chansons de geste, voir Zink 2012.
11. La présente étude étend les observations à toutes les structures propositionnelles annotées dans SRCMF, voir <srcmf.org> pour la taille des extraits.
12. Clari compte 3298 propositions en narration, 437 en OR, Graal 3199 en narration, 2248 en OR et Roland 2471 en narration, 1720 en OR. Clari contient ainsi nettement moins de discours direct. Le fait de regrouper *a priori* les deux informations permet cependant de traiter de la même manière les textes présentant beaucoup d'OR et ceux ayant peu d'OR et de lisser la différence. Cela permet ainsi d'éviter que l'OR soit associé à un type de texte uniquement parce que ce texte présente plus d'OR.
13. Voir Bischof et al. 2012 pour une définition des structures syntaxiques selon le modèle SRCMF.
14. Les individus sont analysés dans un tableau tel que le tableau 1, comprenant 10 colonnes correspondant aux variables ci-dessus, une ligne correspondant à une proposition. Chaque ligne de chaque colonne peut être renseignée par l'une des modalités définie pour chaque variable, ce qui fait ainsi en théorie 92 160 profils de 'ligne' possible, dont seulement 1987 sont attestées.
15. Les interférences suivantes sont évidentes : (i) les formes verbales à l'impératif n'ont pas de sujet exprimé ; (ii) les objets directs et les attributs du sujet s'excluent mutuellement, de même que les objets directs et les sujets impersonnels ; (iii) les subordonnées sont extrêmement rarement paratactiques ; (iv) les relatives ont trop de sujets et d'objets exprimés (par le pronom) pour que cette modalité puisse être prise en compte ; (v) les phrases sans négation *ne* n'ont pas de

particule de négation. Concrètement, les profils de ces modalités ont été remplacés par le profil moyen de la variable.

16. Retirer ces subordonnées de l'échantillon ne change rien au plan factoriel : la disposition des autres modalités reste *grosso modo* la même.

RÉSUMÉS

Dans cette contribution, les auteurs interrogent la stabilité des spécificités syntaxiques de l'oral représenté à travers trois textes de genres distincts en ancien français (chanson de geste, roman en prose et chronique en prose). La distinction entre narration et oral représenté a un effet relativement stable pour la période. Cependant, d'autres facteurs ont aussi un impact considérable sur les structures syntaxiques employées et doivent être pris en compte : il est surtout important de distinguer les propositions principales des subordonnées.

This paper investigates the stability of syntactic characteristics of direct speech (Fr. "oral représenté"). Through the statistical analysis of three texts of different types, it shows that the contrasted distribution of syntactic structures between narrative sentences and direct speech is stable. Nevertheless, other factors also have a crucial impact on syntactic structures. It is the distinction between sentences and subordinate clauses that displays the greatest influence.

INDEX

Keywords : Old French, syntax, lexical statistics, MCA, chanson de geste

Mots-clés : Ancien français, syntaxe, statistiques textuelles, Analyse factorielle des Correspondances Multiples, chanson de geste

AUTEURS

NICOLAS MAZZIOTTA

U. Liège & U. Stuttgart

JULIE GLIKMAN

U. Strasbourg, LiLPa, FDT